**AI - Planet Task**

**Dataset description:**

**Data Source**

\* 43,000+ patient inquiries from real-life scenarios

\* Collected from 12 National Institutes of Health (NIH) websites, including cancer.gov, niddk.nih.gov, GARD, and MedlinePlus Health Topics

**Content**

\* Paired medical questions and answers:

\* Questions cover 37 distinct types, such as treatment, diagnosis, side effects, prevention, and symptoms.

\* Topics range across diseases, drugs, medical tests, and other healthcare entities.

\* Answers provided by healthcare professionals:

\* Doctors

\* Nurses

\* Pharmacists

**Structure:**

Organized into columns:

qtype: Categorizes the question type (e.g., treatment, side effect)

Question: Actual text of the patient's question

Answer: Expert response to the question

Dataset Resource: [Comprehensive Medical Q&A Dataset](https://www.kaggle.com/datasets/thedevastator/comprehensive-medical-q-a-dataset)

**Model**:

Base Model: MiniChat 1.5-3B ([GeneZC/MiniChat-1.5-3B · Hugging Face](https://huggingface.co/GeneZC/MiniChat-1.5-3B))

Finetuned model: <https://huggingface.co/rajveer43/MiniMedicXpert>

Model structure

PeftModelForCausalLM(

(base\_model): LoraModel(

(model): LlamaForCausalLM(

(model): LlamaModel(

(embed\_tokens): Embedding(49216, 3072, padding\_idx=0)

(layers): ModuleList(

(0-23): 24 x LlamaDecoderLayer(

(self\_attn): LlamaAttention(

(q\_proj): lora.Linear4bit(

(base\_layer): Linear4bit(in\_features=3072, out\_features=3072, bias=False)

(lora\_dropout): ModuleDict(

(default): Dropout(p=0.05, inplace=False)

)

(lora\_A): ModuleDict(

(default): Linear(in\_features=3072, out\_features=16, bias=False)

)

(lora\_B): ModuleDict(

(default): Linear(in\_features=16, out\_features=3072, bias=False)

)

(lora\_embedding\_A): ParameterDict()

(lora\_embedding\_B): ParameterDict()

)

(k\_proj): Linear4bit(in\_features=3072, out\_features=3072, bias=False)

(v\_proj): lora.Linear4bit(

(base\_layer): Linear4bit(in\_features=3072, out\_features=3072, bias=False)

(lora\_dropout): ModuleDict(

(default): Dropout(p=0.05, inplace=False)

)

(lora\_A): ModuleDict(

(default): Linear(in\_features=3072, out\_features=16, bias=False)

)

(lora\_B): ModuleDict(

(default): Linear(in\_features=16, out\_features=3072, bias=False)

)

(lora\_embedding\_A): ParameterDict()

(lora\_embedding\_B): ParameterDict()

)

(o\_proj): Linear4bit(in\_features=3072, out\_features=3072, bias=False)

(rotary\_emb): LlamaRotaryEmbedding()

)

(mlp): LlamaMLP(

(gate\_proj): Linear4bit(in\_features=3072, out\_features=8192, bias=False)

(up\_proj): Linear4bit(in\_features=3072, out\_features=8192, bias=False)

(down\_proj): Linear4bit(in\_features=8192, out\_features=3072, bias=False)

(act\_fn): SiLU()

)

(input\_layernorm): LlamaRMSNorm()

(post\_attention\_layernorm): LlamaRMSNorm()

)

)

(norm): LlamaRMSNorm()

)

(lm\_head): Linear(in\_features=3072, out\_features=49216, bias=False)

)

)

)

Output:

